Journal of Statistical and Econometric Methods, Vol. 9, No. 1, 2020, 89-106
ISSN: 2241-0384 (print), 2241-0376 (online)
Scientific Press International Limited

Compositional Data Analysis – Coherent Forecasting Mortality Model with Cohort Effect

Amos BATIONO[1], Leo ODONGO[2] and Karim DERRA[3]

Abstract

In this paper, an extension of the Coherent forecasts of mortality with compositional data analysis (CoDa) model of Bergeron-Boucher et al. (2017) to cohort effect is proposed applied to data from six African countries. The process of fitting this model starts by adapting the Renshaw and Haberman (2006) to compositional data analysis (CODA) as suggested by Bergeron-Boucher et al. (2017). The proposed CoDa-cohort model generally fits the data better than the original cohort model of Renshaw and Haberman (2006). To get the full CoDa-cohort-coherent model the multiple population factor is included in CoDa-cohort model. Then a comparison between CoDa -coherent and CoDa-cohort-coherent models revealed that they have similar accuracy for the selected countries in West Africa but not for countries in East Africa based on Aitchinson distance (AD). But for merged populations like male and female, the new model, CoDa-cohort-coherent, has generally better fits for Kenya mortality data.

Keywords: Mortality, Compositional data analysis, coda, Coherent, Cohort, Forecast

---

[1] Pan African University, Institute of Basic Science Technology and Innovation, Nairobi, Kenya
[2] Department of Mathematics and Actuarial Science, Kenyatta University, Nairobi, Kenya
[3] Clinical Research Unit of Nanoro, Institute of Research in Health Science, Burkina Faso

Article Info: *Received:* September 7, 2019. *Revised:* October 2, 2019.
Published online: January 5, 2020.

1. Introduction

Mortality forecasting has become almost unavoidable in many fields such as financial, actuarial and health institutions. A number model exists in this field. But the most famous belongs to the group of extrapolative forecasting models, the Lee and Carter (1992) model which is widely used and many other models are based on it due to its accuracy and simplicity. However, However, its limitations have led many studies to propose extensions to overcome some flaws and to adapt it to some specific situations or data. This article is focused on extending the CoDa-coherent mortality forecasting model, presented by Bergeron-Boucher et al. (2017), by incorporating an age-period-cohort effect.

2. Literature Review

2.1 Lee and Carter extensions

Lee and Li (2005) proposed an extension to forecast the mortality of many populations simultaneously and coherently. Indeed, the original Lee-Carter model was meant to forecast a single population mortality rate and so it could not take into account two populations male and female simultaneously, for instance, while forecasting. When the forecasts of populations do not diverge over the long run, it is said to be coherent. To avoid that divergence, Lee and Li (2005) used the notion of common factor that captures the effect induced by the common characteristic of all the populations on the rate of mortality. This common characteristic is their belonging to the same group.

Sometimes mortality rates are affected by events that occurred in a specific year. This effect is usually captured considering the year of birth as a cohort. Renshaw and Haberman (2006) added an age–period–cohort term to the original Lee-Carter model to get its cohort-based version.

2.2 CoDa Models

In 2008 Oeppen adapted Lee and Cartel's original model to compositional data. Compositional data can be seen as vectors of components those values are strictly positive with sum equal to a constant. The space of such kind of data is called simplex space. For that, standard statistical analysis is affected (Aitchison 1986, Bergeron-Boucher et al., 2017). Since life table death distribution is under the same constraint, Oeppen (2008) proposed to treat using CoDa framework. This framework provides tools especially for these kinds of data. Oeppen (2008) used the centred log-ratio transformation to bring the data out of the simplex space to the real space before starting analysing them. The advantage of this transformation is that it conserves distance from the simplex to Euclidean spaces.

Recently Bergeron-Boucher et al. (2017) proposed a coherent forecasting model for compositional data. The authors extended the Oeppen (2008) model to coherent forecasting by using the notion of common factor used in Lee and Li (2005). CoDa model solves some problems of Lee and Carter's original model such as fixed rate of mortality improvement (RMI) (Bergeron-Boucher et al., 2017). Indeed, many studies have shown that the assumption of fixed RMI is not supported by empirical studies (Yue et al., 2008). The CoDa coherent model has some advantages but it does not take into account the cohort effect which when ignored can lead to misleading results.

This study is focused on including a cohort effect into Bergeron-Boucher et al. (2017) model. In the first part of this article, we adapt the model of Renshaw and Haberman (2006) to CoDa. In the second part, we will include the cohort effect in the CoDa-coherent model of Bergeron-Boucher et al. (2017); finally, the models will be applied to real mortality data of some African countries. One of the main advantages these CoDa models is that, unlike the original Lee and Carter model, they allow the rate of mortality improvement to change over time.

3. Main Results

We proposed CoDa-cohort model using Renshaw and Haberman (2006) method and we applied it to real Life Tables of six African countries from the World Health Organisation (WHO) website. Compared to the original Renshaw and Haberman (2006) cohort model, the CoDa-cohort model has produced the lowest AIC and BIC for male mortality data of the six countries. Then, we applied the new CoDa-cohort-coherent model presented in this article to the same data of the six African and male and female data of two of these countries. For separated populations like countries, the CoDa-cohort-coherent did not bring additional information (low Aitchinson distance) than the simple CoDa-coherent. But for merged populations like males and females, CoDa-cohort-coherent model brought additional information for Kenya mortality data.

4. Methodology

In this document, $\boldsymbol{x}$ represent the ages; they can single ages (i.e. 18, 20, 30 …) or group of ages (i.e. [1,4], [5,9], …). In any case, their index varies from $\mathbf{1}$ to $\boldsymbol{k}$. The years are represented by $\boldsymbol{t}$ and their index varies from $\mathbf{1}$ to $\boldsymbol{n}$.

4.1 Lee and Carter (1992) and Oeppen (2008) models

4.1.1 Lee and Carter (1992)

The original Lee and Carter (1992) model is presented as follows:

$$\log m_{t,x} = \alpha_x + k_t \beta_x + \epsilon_{t,x} \qquad (1)$$

where m_{tx} represent the central death rate at age x and year t, α_x the average age-specific mortality at age x estimated by the average of $\log m_{t,x}$ for each age over time. The terms k_t and β_x represent respectively the mortality index at the year t and the deviation in mortality rate due to changes in the k_t and are estimated using singular value decomposition (SVD) on the matrix

$$\log m_{t,x} - \alpha_x = k_t \beta_x + \epsilon_{t,x} \qquad (2)$$

$\epsilon_{t,x}$ are the estimation errors with 0 mean and variance σ^2. To determine a unique solution of equation (1), Lee and Carter introduced some constraints to their model

$$\sum_x \beta_x = 1 \quad \text{and} \quad \sum_t k_t = 0 \qquad (3)$$

4.1.2 Oeppen (2008) CoDa model

In the field of forecasting mortality using compositional analysis, Oeppen (2008) proposed a model as an adaptation of the original Lee-Carter model. This CoDa version uses the number of deaths d_{tx} directly instead of the rate of mortality m_{tx}, and the centred log-ratio transformation defined as follows:

$$clr(d_{t,x}) = log\left(\frac{d_{t,x}}{g_t}\right)$$

where g_t is the geometric mean of year t of the numerator, Oeppen (2008) proposed the following formulation

$$clr(d_{t,x} \ominus \alpha_x) = k_t \beta_x + \epsilon_{t,x} \qquad (4)$$

where α_x is the age-specific geometric mean of d_{tx} over time, k_t and β_x are respectively the time index and the age pattern due to a change in k_t. They are found by using SVD on equation (4).

$$k_t = U_t S \text{ and } \beta_x = V_x$$

See Oeppen (2008) for more details.
After fitting the model, d_{tx} is written back to the simplex space using the inverse of centered log-ratio clr^{-1} on equation (4). Then, the perturbation $\oplus$ by α_x gives

$$d_{t,x} = \alpha_x \oplus C[e^{k_t \beta_x + \epsilon_{t,x}}] \qquad (5)$$

The operator $\ominus$ is a perturbation operator of CoDa framework. It is used to centre the matrix without changing the constant sum, the operator $\oplus$ is used to reverse the operation done by $\ominus$; and $C[...]$ is also a CoDa operator called closure, which is used to transform the data into compositional data. All these operators are well described in Bergeron-Boucher et al. (2017).
As in the original Lee and Carter model, forecasts are made by fitting k_t using a time series method.

4.2 Coherent forecasting

4.2.1 Lee and Carter - coherent

To forecast mortality of different populations in a group coherently, that is to avoid divergence in the long run, Li and Lee (2005) proposed an extended model by using a common and specific factor in the original model proposed by Lee and Carter (1992). Their model has been formulated as follows:

$$\log m_{t,x,i} = \alpha_{x,i} + k_t\beta_x + k_{t,i}\beta_{x,i} + \epsilon_{t,x,i} \tag{6}$$

where $m_{t,x,i}$ is the mortality rate at age x and year t of the i^{th} population, $\alpha_{x,i}$ represents the average age-specific mortality at age x in the population i estimated by the average of $\log m_{t,x,i}$ for population i at age x over time, $k_t\beta_x$ is the common factor found by applying the Lee and Carter(1992) to the average mortality of the whole group. The term $k_{t,i}\beta_{x,i}$ is the same term as in Lee and Carter model but specific to each population i and $\epsilon_{t,x,i}$ is the error term specific to the population i.

Li and Lee (2005) noticed that $k_{t,i}$ must approach to a constant over time for the model to work (that is to avoid divergence in mortality rate for each population i of the same group over the long run). They also noticed that the model does not guarantee this convergence to a constant and that can make the model fail. Therefore, they suggest to use a random walk without drift or a first-order autoregressive model with a constant term (Bergeron-Boucher et al., 2017).

4.2.2 CoDa- coherent

Following the same logic to build a coherent mortality forecasting model for CoDa, Bergeron-Boucher et al. (2017) modified the model proposed by Oeppen (2008). They also considered a common and specific factor for the CoDa model of Oeppen (2008) as Li and Lee (2005) did for the original model proposed by Lee and Carter (1992). The CoDa-coherent model proposed by Bergeron-Boucher et al. (2017) is given by

$$clr\left(d_{t,x,i} \ominus \alpha_{x,i} \ominus C[e^{k_t\beta_x}]\right) = k_{t,i}\beta_{x,i} + \epsilon_{t,x,i} \tag{7}$$

It is therefore similar to the Lee and Carter coherent model proposed by Li and Lee (2005). Here also, $\alpha_{x,i}$ is the geometric mean of the number of deaths at age x in the population i. The term $C[e^{k_t\beta_x}]$ is common for every population of the group since it does not depend on i. It is found by applying Oeppen (2008) CoDa model to the average $d_{t,x}$ of the whole group.

The term $k_{t,i}\beta_{x,i}$ represents the perturbation factor specific to every population i and is obtained by applying SVD to equation (7). The time series $k_{t,i}$ is also forecasted using time series methods that make it approach to a constant. As for Li and Lee (2005), the authors the CoDa-coherent model proposed does not guarantee that $k_{t,i}$ will convergence.

4.3 Cohort effect models

4.3.1 Lee and Carter – cohort model

To consider the cohort effect in mortality forecasting, Renshaw and Haberman (2006) extended the model proposed by Lee and Carter (1992) by adding a supplementary factor $l_{t-x}\beta_x^{(0)}$. Their extended Lee and Carter mortality model with cohort effect is then given by

$$\log m_{t,x} = \alpha_x + l_{t-x}\beta_x^{(0)} + k_t\beta_x^{(1)} + \epsilon_{t,x} \tag{8}$$

The term $l_{t-x}\beta_x^{(0)}$ represents the cohort effect. Since t is the year and x the age, the year of birth is then represented by $t-x$.

As for Renshaw and Haberman (2006), Oeppen (2008), modelling will be done directly on the number of deaths ($d_{t,x}$) instead of on the death rate ($m_{t,x}$) proposed by Lee and Carter (1992). Such modelling is useful if we want to simulate future cash flows of an annuity for instance (Renshaw and Haberman, 2006). In order to identify a unique solution, the authors imposed the following constraints

$$\sum_x \beta_x^{(0)} = 1 \qquad \sum_x \beta_x^{(1)} = 1 \qquad \text{and} \qquad \sum_t k_t = 0$$

4.3.2 CoDa-cohort

Similarly, as Oeppen (2008) did for Lee and Carter model by using CoDa framework, we propose to write the CoDa version of Renshaw and Haberman (2006) cohort model as below with a little difference

$$y_{t,x} = \frac{d_{t,x}}{g_t} = \alpha_x \oplus C[e^{l_{t-x}\beta_x^{(0)} + k_t\beta_x^{(1)} + \epsilon_{t,x}}] \tag{10}$$

Where g_t is the geometric mean of $d_{t,x}$ of year t; α_x is the average age-specific mortality (centred other time). The term $k_t\beta_x^{(1)}$ is the same as for the original Lee and Carter (1992) model and $l_{t-x}\beta_x^{(0)}$ represents the cohort effect. Depending on the values assigned to $\beta_x^{(0)}$ and $\beta_x^{(1)}$ it is possible to define sub-categories of the model in (10) similar to those in Renshaw and Haberman (2006).

Considering the expected value part, we can re-write $y_{t,x}$ as

$$\frac{d_{t,x}}{g_t} = e^{\log \alpha_x} \oplus C[e^{l_{t-x}\beta_x^{(0)} + k_t\beta_x^{(1)}}]$$

$$\frac{d_{t,x}}{g_t} = C[e^{\log \alpha_x + l_{t-x}\beta_x^{(0)} + k_t\beta_x^{(1)}}] \tag{11}$$

And using the logarithm transformation we have

$$\log \frac{d_{t,x}}{g_t} = clr(d_{t,x}) = \log \alpha_x + l_{t-x}\beta_x^{(0)} + k_t\beta_x^{(1)} \tag{12}$$

Therefore, since $y_{t,x} = \frac{d_{t,x}}{g_t}$, from (10) we can write

$$\log y_{t,x} = \lambda_x + l_{t-x}\beta_x^{(0)} + k_t\beta_x^{(1)} \tag{13}$$

(where $\lambda_x = \log \alpha_x$).
Which is the corresponding link function in Renshaw and Haberman (2006) under Poisson assumption.
Since the centred log-ratio (clr) transformation brings data out of simplex space to the Euclidean space, the parameters in (13) can be estimated using standard methods such as the iterative procedure of deviance minimization used by Renshaw and Haberman (2006). Under Poisson assumption, the deviance $D(.,.)$ is given by

$$D(y_{t,x}, \hat{y}_{t,x}) = \sum_{t,x} 2w_{t,x}\left(y_{t,x}\log\left(\frac{y_{t,x}}{\hat{y}_{t,x}}\right) - (y_{t,x} - \hat{y}_{t,x})\right) \tag{14}$$

Since our main variable of interest is $d_{t,x}$ instead of $y_{t,x}$, we will not minimize the deviance of $y_{t,x}$ but that of $d_{t,x}$. Therefore, $y_{t,x}$ and $\hat{y}_{t,x}$ must be replaced respectively by $d_{t,x}$ and $\hat{d}_{t,x}$ in equation (14).

Considering different geometric mean g_t for each year leads to high fluctuations and poor results. This problem is fixed when we consider that every year has the same general geometric mean which is their average over the years. Therefore g_t will be taken as the average for all years in the equation

$$\frac{d_{t,x}}{g_t} = y_{t,x}$$

During the iterative procedure, the parameters are updated using Newton-Raphson deviance minimisation method. In general, a parameter θ is updated using the formula

$$u(\theta) = \theta - \frac{\frac{\partial D}{\partial \theta}}{\frac{\partial D^2}{\partial \theta^2}}$$

And under Poisson assumption, parameters updating formula are given as follows

$u(\hat{\lambda}_x)$	$=$	$\hat{\lambda}_x + \frac{\sum_t w_{t,x}(y_{t,x} - \hat{y}_{t,x})}{\sum_t w_{t,x}\hat{y}_{t,x}}$
$u(\hat{l}_z)$	$=$	$\hat{l}_z + \frac{\sum_{t-x=z}^{t,x} w_{t,x}\left(y_{t,x} - \hat{y}_{t,x}\right)\hat{\beta}_x^{(0)}}{\sum_t w_{t,x}\hat{y}_{t,x}\,\hat{\beta}_x^{(0)^2}}$
$u(\hat{\beta}_x^{(0)})$	$=$	$\hat{\beta}_x^{(0)} + \frac{\sum_t w_{t,x}(y_{t,x} - \hat{y}_{t,x})\hat{l}_{t-x}}{\sum_t w_{t,x}\hat{y}_{t,x}\,\hat{l}_{t-x}^2}$
$u(\hat{k}_t)$	$=$	$\hat{k}_t + \frac{\sum_t w_{t,x}(y_{t,x} - \hat{y}_{t,x})\hat{\beta}_x^{(1)}}{\sum_t w_{t,x}\hat{y}_{t,x}\,\hat{\beta}_x^{(1)}}$
$u(\hat{\beta}_x^{(1)})$	$=$	$\hat{\beta}_x^{(1)} + \frac{\sum_t w_{t,x}(y_{t,x} - \hat{y}_{t,x})\hat{k}_{t-x}}{\sum_t w_{t,x}\hat{y}_{t,x}\hat{k}_{t-x}^2}$

where $w_{t,x}$ is equal to 1 if $d_{t,x} > 0$ and 0 if $d_{t,x} = 0$.
(see Renshaw and Haberman, 2006).

The iterative algorithm for fitting the CoDa-cohort model is similar to the one proposed by Renshaw and Haberman (2006), is given as follows:

Step 1: Set starting values

- $\hat{\lambda}_{\mathrm{x}} = \left(\prod_t y_{t,x}\right)^{1/n}$
- $\hat{\beta}_x^{(0)} = 1$
- $\hat{\beta}_x^{(1)} = \frac{1}{k}$
- $\hat{k}_t = \hat{l}_{t-x} = 0$

Step 2: Update $\hat{\lambda}_x$

Step 3: Given $\hat{\beta}_x^{(0)}$ and $\hat{\beta}_x^{(1)}$, update $\hat{k}_t$ and $\hat{l}_{t-x}$

Step 4: Given $\hat{k}_t$ and $\hat{l}_{t-x}$, update $\hat{\beta}_x^{(0)}$ and $\hat{\beta}_x^{(1)}$

Step 5: Compute $D(d_{t,x}, \hat{d}_{t,x})$

where k is the total number of ages (or groups of ages).

As in Renshaw and Haberman (2006), we repeat the cycle from step 2 until $D(d_{t,x}, \hat{d}_{t,x})$ converges ($\Delta D \approx 0$).
As for most of the mortality forecasting models based on Lee and Carter (1992), forecasting the number of deaths is based on time series models. Only the parameters involving time (i.e. $\hat{k}_t$ and $\hat{l}_{t-x}$,) will be forecasted. The predicted value of $d_{t,x}$ is given as follows:

$$\dot{d}_{t_n+s,x} = \hat{\alpha}_x \oplus \left(g.C\left[e^{\dot{l}_{t_n-x+s}\hat{\beta}_x^{(0)}+\dot{k}_{t_n+s}\hat{\beta}_x^{(1)}}\right]\right) \tag{15}$$

The following notations can be found in Renshaw and Haberman (2006)

$$\{\hat{l}_z : z \in [t_1 - x_k, t_n - x_1]\} \mapsto \{\dot{l}_{t_n-x+s} : s > 0\}$$

And

$$\{\hat{k}_t : t \in [t_1, t_n]\} \mapsto \{\dot{k}_{t_n+s} : s > 0\}$$

If $0 < s \leq \mathrm{x} - \mathrm{x}_1$, then $\dot{l}_{t_n-x+s}$ takes the value of estimated value $\hat{l}_{t_n-x+s}$.
The dot on the top of parameters stands for predicted value.
The time series, k_t and l_z, are forecasted using random walk with drift.

4.3.3 CoDa-cohort-coherent

To forecast multiple population mortality coherently, Lee and Li (2005) and Bergeron-Boucher et al. (2017) have both used a common factor in their model to capture the general pattern of mortality of the whole group of populations. In the CoDa-cohort-coherent model presented in this article, as a common factor, we suggest to use, in addition to the general pattern used by Lee and Li (2005) and Bergeron-Boucher et al. (2017), the cohort effect. The model is presented as following

$$d_{t,x,i} = \alpha_{x,i} \oplus g.C\left[e^{l_{t-x}\beta_x^{(0)}+k_t\beta_x^{(1)}}\right] \oplus C\left[e^{k_{t,i}\beta_{x,i}+\epsilon_{t,x,i}}\right] \tag{16}$$

Where $\alpha_{x,i}$ is the age-specific geometric mean of $d_{t,x,i}$, all the parameter that does not involve i form together the common factor and they are found by applying the CoDa-Cohort method to the average $d_{t,x}$ of the whole group. The errors are represented by $\epsilon_{t,x,i}$. The remaining parameters, $k_{t,i}$ and $\beta_{x,i}$, specific to each population i, are found by applying SVD to the following matrix as Bergeron-Boucher et al. (2017) did for CoDa-coherent model

$$clr\left(d_{t,x,i} \ominus \alpha_{x,i} \ominus g.C\left[e^{l_{t-x}\beta_x^{(0)}+k_t\beta_x^{(1)}}\right]\right) = k_{t,i}\beta_{x,i} + \epsilon_{t,x,i} \tag{17}$$

As for the previous coherent forecasting models, the time series $k_{t,i}$ is also forecasted using time series methods that make $k_{t,i}$ approach a constant and there is no guarantee that $k_{t,i}$ will reach that convergence.

5. Results

5.1 Data

The data used in this study come from the World Health Organization (WHO) website: http://apps.who.int/gho/data/node.main.LIFECOUNTRY?lang=en). The web site provides life table by country from 2000 to 2016. We will choose for our study mortality data from six African countries: Burkina Faso, Ivory Coast and Mali from West Africa; Ethiopia, Kenya and Somalia from Eastern Africa. Two situations will be considered: The first will be to consider only male mortality data from these six countries and the second, male and female mortality data of two countries, Burkina Faso and Kenya.

The total number of deaths for some years does not reach 100,000 (the table radius). For such cases, the remaining number of deaths for each year will be added to the number of deaths of the age with the higher number of deaths in that year.

5.2 Cohort models comparison

In this section, we will compare the CoDa cohort model with that of Renshaw and Haberman (2006) model (RH cohort). We will apply those two models to each of the six countries and compare their goodness of fit using their Akaike information criterion (AIC) and Bayesian information criterion (BIC) based on the observed $d_{t,x}$ and their respective fitted value. Table 1 contains the AIC and the BIC.

Compared to the RH cohort model, and using the AIC and BIC as a model selection criterion, we can notice that the CoDa cohort has the lowest value of AIC and BIC for all the six countries. Therefore, we can say that the CoDa cohort model fits these data better than the Renshaw and Haberman (2006) cohort model.

Table 1: Cohort models Akaike information criterion (AIC) and Bayesian information criterion (BIC)

	RH cohort		**CoDa cohort**	
Country	**AIC**	**BIC**	**AIC**	**BIC**
Burkina	3447.254	3826.903	3360.036	3739.685
Ivory Coast	3604.273	3983.922	3554.101	3933.75
Mali	3461.602	3841.252	3380.949	3760.598
Kenya	3695.329	4074.978	3469.57	3849.219
Somalia	3575.096	3954.745	3468.503	3848.152
Ethiopia	3546.997	3926.646	3460.816	3840.465

Source: WHO (2000 – 2016) – Authors calculations

5.3 CoDa coherent models

In this section, men mortality data for the six countries has been fitted using CoDa-coherent model Bergeron-Boucher et al. (2017) and the CoDa-cohort-coherent presented in this study. We considered two groups of countries based on their proximity criteria and their region in Africa. Figure 1 represents the patterns of β_{xi} by ages and Table 2 contains the Aitchinson distance (AD) to measure the dissimilarity in compositional data analysis (see Bergeron-Boucher et al. (2017) or Aitchison et al. (2000) for more details)

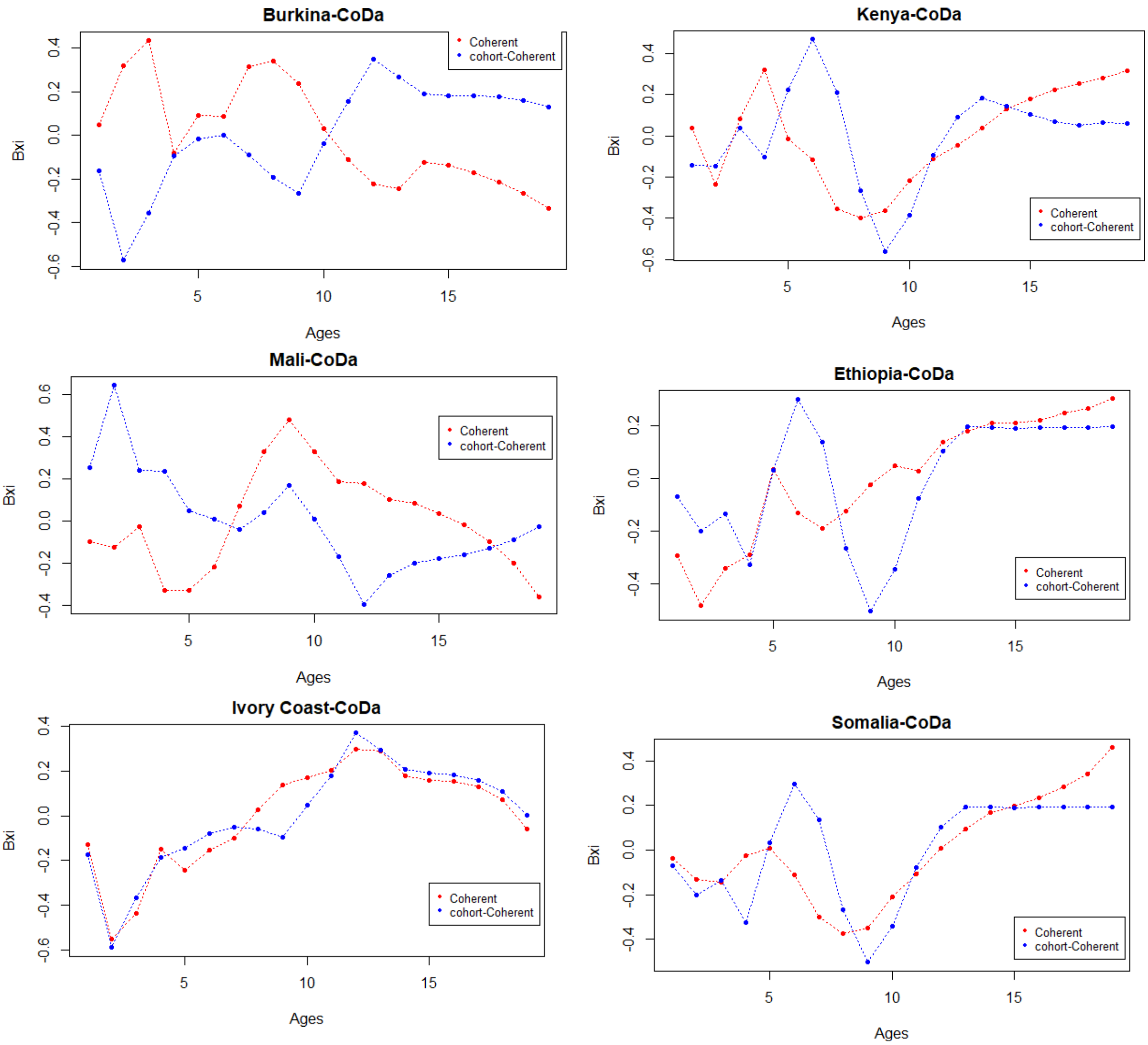
$$AD = \left[\sum_{\forall i}\left(clr(x_i) - clr(y_i)\right)^2\right]^{1/2}$$

where x_i is observed composition and y_i the fitted composition.
Table 2 also contains the mean error (ME) which will be used here to measure the fitting bias of the two methods on the two groups of countries (See Bergeron-Boucher et al. (2017) for details).

$$ME = \left[e_0^{Expected} - e_0^{Observed}\right]$$

Considering the pattern of β_{xi} we have noticed that the highest value has been recorded for CoDa-cohort-coherent and simple CoDa-coherent in the group of West Africa (Mali for people from 1 to 4 years old) and in the group of East Africa (Somalia for people of more than 85 years old) respectively. For higher ages, all the two coherent models show an increasing value in the patterns of β_{xi} for East Africa countries and a decreasing value for West Africa countries except Mali for the CoDa-cohort-coherent. For Ivory Coast, the two models have produced almost the same values.

West Africa **East Africa**



Ages: 5 is people aged 15 to 19 years, 10 is 40 to 44 years old and 65 to 69 years old.
Source: WHO (2000 – 2016) – Authors calculations

Figure 1: Pattern β_{xi} specific to the countries of each of the two groups: West Africa and East Africa

Considering Aitchinson's distance, we notice that the models have similar AD for the West Africa countries. Although considering the cohort effect, CoDa-cohort-coherent model fits these data of West Africa similarly as the simple CoDa-coherent model does. They also have the same level of bias in their fitting methodology since the ME both of the CoDa-coherent and CoDa-cohort-coherent have the same number of positive values. But for the group of East Africa countries, the CoDa-cohort-coherent model produced the higher value of AD with negative ME. So, including cohort effect in the effect for these countries did not bring significant additional information. A possible explanation is that the data of these countries, together, do not share significant common cohort effect.

Table 2: Comparison of CoDa coherent and CoDa cohort coherent models based on ME and AD

	CoDa Coherent		**CoDa Cohort Coherent**	
Country	**ME**	**AD**	**ME**	**AD**
Burkina	0.0007152	14.14071	0.000712	14.43345
Ivory Coast	0.0004034	11.28661	0.000402	11.65467
Mali	0.0007239	14.00225	0.000717	14.27725
Kenya	0.0006271	13.87013	-0.000942	30.61011
Somalia	0.0005142	12.99342	-0.001122	30.45637
Ethiopia	0.0006622	13.42291	-0.000915	30.12357

Source: WHO (2000 – 2016) – Authors calculations

Now we consider another situation. We select two countries (Burkina Faso and Kenya) and apply CoDa-coherent and CoDa-cohort-coherent on male and female mortality data of these two countries. So, we have two groups (the countries) and each group has two populations (male and female). Table 3 contains the AD and ME of these two coherent models.

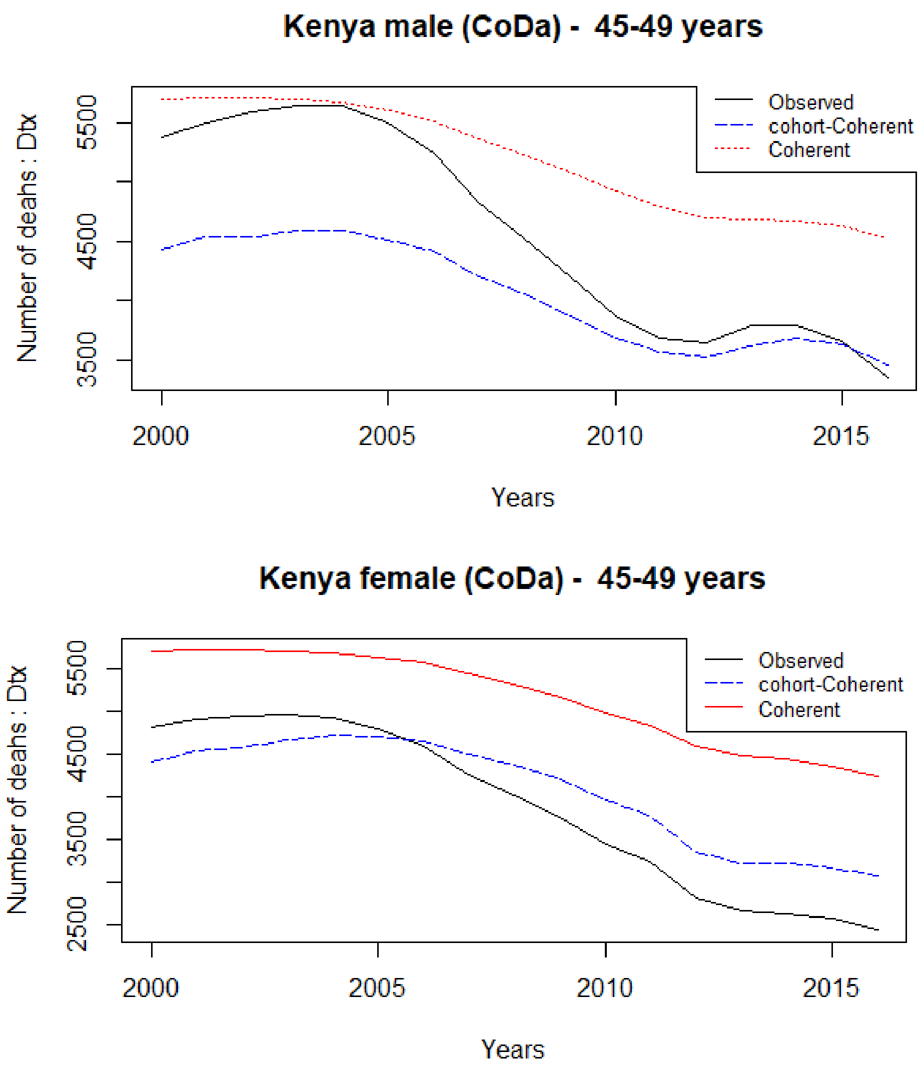
In terms of bias, none of the two models has a negative value of ME. Therefore, any of them has more bias than the other. Considering AD, we notice that including the cohort effect in the mortality model does not bring additional information compared to simple CoDa-coherent model for Burkina Faso. But as far as Kenya is concerned, considering the cohort effect bring an additional information since the CoDa-cohort-coherent model has the lowest AD, and then, produces a better fit than the CoDa-coherent model of Bergeron-Boucher et al. (2017).

Table 3: Comparison of CoDa coherent and CoDa cohort coherent models based on ME and AD

		CoDa Coherent		CoDa Cohort Coherent	
Country		**ME**	**AD**	**ME**	**AD**
Burkina	Male	0.0007152	14.1348	0.0006778	14.34837
	Female	0.000811	14.06167	0.0007782	14.26234
Kenya	Male	0.0006251	13.86435	0.0005566	12.93618
	Female	0.0009624	16.07809	0.000909	15.21048

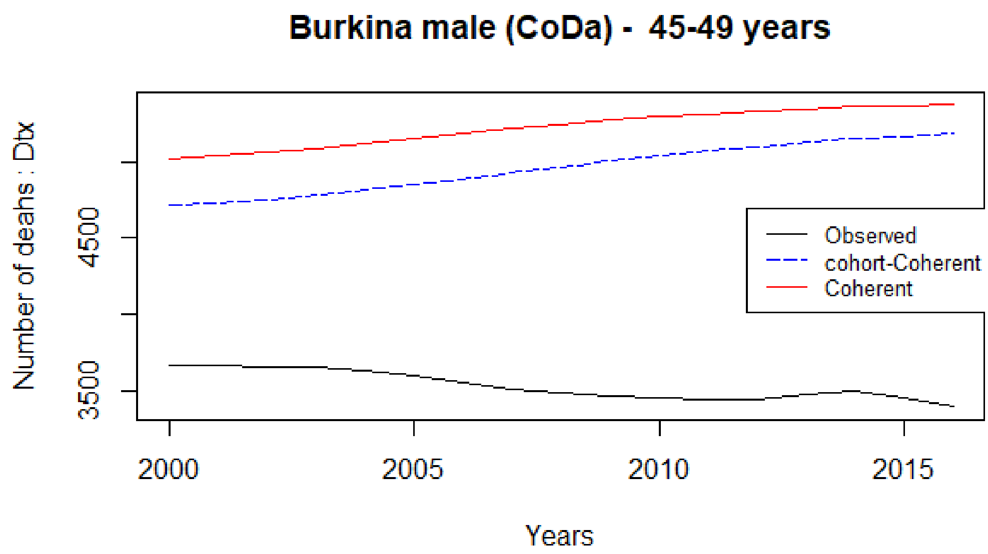
Source: WHO (2000 – 2016) – Authors calculations

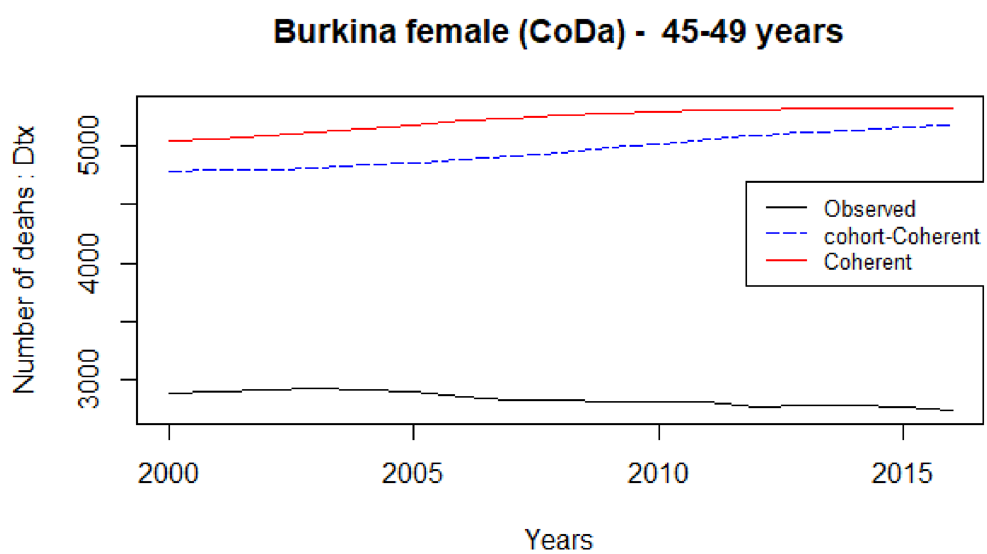
Let's consider fitted data of the two coherent models. Figure 2 and Figure 3 contain pattern of fitted and observed $d_{t,x,i}$ by sex over years of people aged 45 to 49 years old (they are neither young nor seniors) for Kenya and Burkina Faso respectively. The black curve is the observed $d_{t,x,i}$ the red curve represents fitted $d_{t,x,i}$ using simple CoDa-coherent and the blue is the one for CoDa-cohort-coherent model. Both coherent models draw the same pattern as the observed $d_{t,x,i}$ for 45 to 49 years old male of the two countries. Both of the fitted data of the two models remain at a distance from the observed data for Burkina Faso. Considering Kenya, the fitted male data of CoDa-coherent model was the closest to the observed male data from 2000 to 2005; but from that year we observed a rapid decreasing in male mortality which the simple CoDa-coherent did not follow since it decreased slowly. Then, the CoDa-cohort-coherent model becomes the closest to the observed data up to the end (2016). For female mortality data, CoDa-cohort-coherent model produced the most accurate fitted data for all the years in Kenya.



Source: WHO (2000 – 2016) – Authors calculations

Figure 2: Fitted $d_{t,x}$ Kenya, Male and Female, 45-49 years old





Source: WHO (2000 – 2016) – Authors calculations

Figure 3: Fitted $d_{t,x}$ Burkina Faso, Male and Female, 45-49 years old

6. Conclusion

The prediction of mortality is a field that continues to grow as the amount of research devoted to it is large. Despite a fairly recent introduction, CoDa is already starting to have a major impact on this area. From the Lee and Carter CoDa version of Oeppen (2008) to CoDa coherent of Bergeron-Boucher et al. (2017), this study continues in the same direction by proposing a Renshaw and Haberman (2006) cohort CoDa version and an extension of Bergeron-Boucher et al. (2017) CoDa coherent model to cohort (CoDa-cohort-coherent).

By using male mortality data of six African countries and AIC and BIC as a model selection criterion, it is observed that the CoDa-cohort model offers a better fit compared to the original RH cohort model. In the case of coherent models, the simple CoDa-coherent and the CoDa-cohort-coherent model, the same level of bias in their fitting method was observed for the group of West Africa countries. But considering the group of East Africa countries, introducing cohort effect has not produced a better fit. This may mean that these countries, taken together, do not share a significant common cohort effect, or some of them are too different in terms of mortality pattern affecting the average mortality to represent an accurate general trend. Considering merged populations as male and female, we noticed that in general, the CoDa-cohort-coherent model has a better fit. Therefore, the CoDa-coherent model of Bergeron-Boucher et al. (2017) will be well used for data without cohort effects. But, in the interest of the forecast is cohort forecasting, CoDa-cohort-coherent must be used but the populations of that group must share significant historical events affecting life expectancy to create a common cohort effect, otherwise, the cohort forecast might fail. It must be noticed that the CoDa-cohort fitting process has been done based on some assumptions.

The first assumption is the value of that has been set to the average of the geometric mean of all years. That means every year has the same geometric mean of the number of deaths. This assumption was necessary to reduce the fluctuation to get a more stable and accurate estimation. The second assumption comes from the formulation of the CoDa-cohort-coherent model (equation 16). Indeed, the age-cohort term does not involve. That means every individual with the same year of birth. It doesn't matter which population he/she comes from, will have the same cohort effect. That might work fine when the populations are merged (for instance male and female in a country, town, etc.). But, if the populations are countries, regions, etc. each of these populations are likely to have cohort effects specific to each of them. Such an approach could have a better fit for a group like the East Africa one for instance. This can be an issue for further research.

ACKNOWLEDGEMENTS. The authors are grateful to the African Union (AU) for their unfailing support, the Pan African University Institute for basic Sciences Technology and Innovation (PAUSTI) for their assistance and the anonymous viewers for their useful comments.

References

[1] A. E. Renshaw, and S. Haberman, A Cohort-Based Extension to the Lee – Carter model for Mortality Reduction Factors. *Insurance: Mathematics and Economics*, **38**, (2006), 556–570.

[2] J. Aitchison, The Statistical Analysis of Compositional Data, *Journal of the Royal Statistical Society. Series B (Methodological),* Vol. 44, No. 2. (1982), pp.139-177.

[3] J. Oeppen, Coherent forecasting of Multiple decrement Life Tables: a test using Japanese cause of death data. (1992). *Paper presented at the European Population Conference, Barcelona, Spain*, July 9–12, 2008

[4] N. Li and R. Lee, Forecasts for a Coherent Mortality an Extension of Group of Population: The Lee-Carter. *Demography*, **42**(3), (2005), 575–594.

[5] M.P. Bergeron-Boucher, V. Canudas-Romo, J. Oeppen, and J. W. Vaupel, Coherent Forecasts of Mortality with Compositional Data Analysis. *Demographic Research*, **37**(August), (2017), 527–566.

[6] R. D. Lee, and L. R. Carter, Modelling and Forecasting U.S. Mortality, *Journal of the American Statistical Association*, (1992)